



PACE™ Technology

The Realization of Multimedia-Enabled Ethernet



More connected.™

PACE™ Technology

The Realization of Multimedia-Enabled Ethernet

Contents

Introduction	2
PACE Interactive Access	3
Description of PACE Interactive Access	4
Performance of PACE Interactive Access for 10 Mbps Link Speed	4
Performance of PACE Interactive Access for 100 Mbps Link Speed	7
PACE Repeater	8
Description of the PACE Repeater Algorithm	8
Performance of PACE Repeater for Video Traffic Alone	9
Performance of PACE Repeater for Data Traffic Alone	11
Performance of PACE Repeater for Integrated Video and Data Services	14
PACE Repeater Scalability	14
Conclusion	16
References	17

Acronyms and Abbreviations

BER

bit error rate

BLAM

Binary Logarithmic Access Method

CRC

cyclic redundancy check

CSMA/CD

Carrier Sense Multiple Access/Collision Detection

FCFS

first come, first served

IEEE

Institute of Electrical and Electronics Engineers

IPG

inter-packet gap

LAN

local area network

MAC

Media Access Control

NIC

network interface card

PE-MAC

PACE technology-enabled MAC

PLR

packet loss rate

PACE Technology

The Realization of Multimedia-Enabled Ethernet

While Ethernet is the most commonly used LAN scheme, it has some shortcomings for supporting real-time and/or multimedia traffic. In particular, due to the "capture effect," it cannot support audio/video traffic in the presence of bursty data traffic even when the network utilization is fairly low (i.e., 10 to 15 percent).

In this paper, we introduce PACE™ technology, which provides a low-cost, backward-compatible adaptation of Ethernet to support real-time and/or multimedia applications. We first describe the PACE technology for dedicated, switched networks, and show some measurement results that indicate that PACE technology can indeed successfully support real-time multimedia applications. We then introduce a generalization of the PACE algorithm, which makes PACE applicable for shared LAN segments by replacing the standard Ethernet repeater with a "PACE Repeater," while not requiring any changes in the end-station network software, network interface, or wiring. By means of computer simulation, we show that PACE Repeater also provides a significant improvement over standard Ethernet when supporting multimedia traffic.

Introduction

The availability of increasingly powerful multimedia personal computers interconnected by local area networks (LANs) has led to many new distributed multimedia applications. Examples of such applications include desktop video and data conferencing, Internet telephony, distance learning, telemedicine, and telescience. Many of those applications involve transmission of audio and video traffic, which has characteristics and requirements that are different from those of traditional data traffic. In particular, due to the high data volume, and in some cases due to the interactivity requirements, the audio and video traffic has to be sent in a streamed fashion and played back at the receiver in real time. Therefore, in order to maintain the continuity of playback,

data generated at the sender must be delivered to the receiver within a bounded time. Any data delayed more than that time is useless, and considered to be lost. Such loss of information leads to quality degradation in received audio and video, and therefore must be kept as low as practically possible.

Today, the most popular local area network type is Ethernet [1] (in both 10 Mbps and 100 Mbps forms), which has not been designed to support streaming multimedia applications. The Carrier Sense Multiple Access with Collision Detection (CSMA/CD) media access control (MAC) protocol used in Ethernet does not guarantee bounded delays, and behaves poorly under heavy load conditions, leading to excessive latency and delay jitters, as well as packet loss due to excessive collisions. Nevertheless, a number of studies have shown that as long as the traffic in the network consists only of audio and video streams (or similarly well-behaved, non-bursty data traffic), a good deal of network capacity can be utilized while providing good support for the audio/video streams (i.e., 30 to 50 percent for audio, 50 to 70 percent for video, depending on various parameters such as the delay and loss constraints, etc.) [2-5]. Therefore, such well-behaved traffic can be supported by providing sufficient bandwidth by means of segmenting the network using Ethernet switches and/or by deploying Fast Ethernet, and by the appropriate deployment of the end stations so that the network utilization in every segment remains below the above-mentioned limits.

However, Ethernet has another, more serious limitation: it cannot distinguish between different traffic types to allow expedited delivery of audio/video streams. Indeed, it has been shown that when bursty data traffic is present in the network (with a burst size as small as 10 KB), even when the total network utilization is as little as 10 to 15 percent, the audio/video performance is severely degraded [5]. This is in particular due to the well-known "capture effect" inherent in the CSMA/CD protocol with the binary exponential backoff algorithm that is used in the Ethernet [6]. Considering that network traffic is often bidirectional, with

a mix of high-bandwidth audio/video streams and bursty client-server transfers, it is clear that bandwidth availability alone is not sufficient to support multimedia traffic. Indeed, in the presence of bursty data traffic, even dedicating a switch port for every end station is not sufficient to adequately support real-time audio/video traffic. Therefore, the detrimental effect of bursty data traffic on audio/video streams must be circumvented by using a media access control protocol that is less sensitive to traffic burstiness, and by providing a mechanism to distinguish between different traffic types.

Over the past decade, a number of attempts have been made to modify the binary exponential backoff algorithm [7], [8] or employ some additional access control [9-13] to enable the Ethernet for real-time and/or multimedia traffic. Most of these efforts have focused on improving the performance of the shared environment. None have resulted in commercial products. Most notable among those is Binary Logarithmic Access Method (BLAM) [8], which has been considered for standardization by IEEE 802.3w. The idea behind BLAM is to provide a backoff mechanism that does not result in as high a delay variance as the binary exponential backoff algorithm. An important advantage of this scheme is that BLAM-enabled end stations can coexist with standard 802.3 Ethernet stations on the same LAN segment. However, since the backoff mechanism is implemented in hardware in the network interface cards (NICs), every BLAM-enabled station must be deployed with a new NIC.

The fundamental difficulty in this field of research is the unpredictable nature of a shared segment where a large number of differing traffic scenarios must be accommodated. In addition, the new mechanism must be backward compatible with the existing nodes that are still running standard binary exponential backoff algorithm to facilitate a practical migration process. Verifying that a new mechanism is backward compatible, that it offers predictable and yet acceptable packet delay and jitter under all circumstances yet imposes minimal overhead so as to maintain

the throughput performance, is indeed a daunting task.

In this paper, we present PACE technology, which addresses this problem. Unlike prior schemes that have been proposed, our main design principle has been to ensure that no changes are required in the end-station NICs, software, or wiring, thus maximizing backwards-compatibility.

We first discuss the original PACE technology that was devised for a switched Ethernet environment with one station connected to any given switch port. It consists of two parts: *PACE Interactive Access* and *PACE Implicit Class of Service*. The first part, PACE Interactive Access, is a modification of the Ethernet binary exponential backoff algorithm at the switch port. PACE Interactive Access provides significantly lower access delay and jitter as compared to the regular Ethernet MAC, as we demonstrate by some traffic measurements. Furthermore, it is fully compatible with the standard Ethernet MAC protocol at the end stations; therefore, it does not require any changes in the end stations. The second part of the PACE technology, PACE Implicit Class of Service, allows delay sensitive traffic to be given higher priority than other traffic in a practical and cost-effective manner. PACE Implicit Class of Service is not covered in this paper¹.

We then introduce a generalization of the PACE Interactive Access technology for the case of a shared LAN segment with multiple end stations, referred to as PACE Repeater. We provide computer simulation results comparing the performance of PACE Repeater with regular Ethernet.

PACE Interactive Access

Recognizing the trend toward switching, PACE Interactive Access was developed as a complementary technology to enable switched Ethernet LANs for real-time and/or multimedia traffic [14]. By taking advantage of the fact that only two stations are on any connection (end station and switch), the switch can be tuned to control the link and guarantee alternate transmissions from itself and the end station. Thus, both jitter and latency are

¹ PACE Implicit CoS was designed to fill in a void until a standard Layer 2 CoS mechanism was devised. With the advent of IEEE 802.1p CoS, it is no longer needed.

bounded well within the limits necessary to support real-time and streaming applications. Additionally, since the PACE Interactive Access algorithm is designed to work with end stations with standard Ethernet NICs, only the MAC silicon in the switch needs to be upgraded. This backward interoperability with existing Ethernet NICs provides the least expensive and most straightforward migration path to enable any Ethernet LAN for networked multimedia applications.

The next section describes the PACE Interactive Access algorithm and gives some measurement results for a 10 Mbps network; showing that the PACE Interactive Access algorithm indeed achieves lower latency and jitter than regular Ethernet. We present results for Fast Ethernet that are similar to those for the 10 Mbps Ethernet, showing that PACE Interactive Access scales well to 100 Mbps.

Description of PACE Interactive Access

PACE Interactive Access is a mode of operation where traffic is ordered by a PACE technology-enabled node to ensure that network jitter and latency are kept to a minimum. Network jitter and latency are combined into a joint parameter known as access latency, defined as the maximum time that a port will take to either successfully transmit a packet or discard it, as measured from the time the packet is presented to the MAC for transmission. For “regular” Ethernet, this access latency can be on the order of seconds (e.g., maximum backoff time for 16 collisions plus time deferring to others using network). Packets with access latency in the order of 300 ms can often be seen on regular Ethernet. To allow multimedia applications to run over Ethernet, PACE Interactive Access has been designed to provide nominal access latency bounded at a 5-10 ms range.

The PACE Interactive Access algorithm works as follows [15]. The PACE technology-enabled MAC (PE-MAC) keeps track of whether or not it was the last port to transmit successfully. When it receives a packet, the PE-MAC moves into “Trying to Transmit” state if it has a packet to send. The PE-MAC will attempt to transmit a packet at the end of

the normal Ethernet inter-packet gap (IPG). If a collision occurs during this transmission, it retries transmission again at the end of another IPG (i.e., backoff time is forced to zero in this state). This aggressive retransmission attempt is repeated for up to (attemptLimit - 1) consecutive collisions. Then, if carrier sense is not detected, the PE-MAC attempts transmission one last time after waiting half a slot-time. If carrier sense is detected at this point, however, no attempt is made to transmit the packet and the packet is discarded. The reason for performing the last attempt in this manner is to guarantee that no collision will occur (thus to ensure that access latency is not increased) while increasing the probability of successfully transmitting the packet. After either transmitting or discarding a packet upon reaching attemptLimit, the PE-MAC moves into “Waiting to Receive” state.

In “Waiting to Receive” state, the PE-MAC defers just long enough to allow the node on the other end of the link a chance to transmit a packet if it has one to send. The backoff time is derived from the number of collisions it has encountered in the previous transmission. In particular, if no collisions are experienced during the previous “Trying to Transmit” state, there will be no extended defer time. In that case, if there is another packet to send, the PE-MAC will attempt to transmit the next packet after the IPG has expired. However, if there were collisions in the previous “Trying to Transmit” state, an extended deference guarantees that the other node will come out of its backoff and transmit its packet.

Upon the expiration of deference timer or receipt of a packet in the “Waiting to Receive” state, the PE-MAC moves back into “Trying to Transmit” state. If the PE-MAC exits the “Waiting to Receive” state without receiving a packet, it will stay in “Trying to Transmit” state and transmit packets separated by the normal Ethernet IPG until a collision is seen.

Performance of PACE Interactive Access for 10 Mbps Link Speed

Table 1 shows the maximum access latency and the statistical packet loss rate as a function

Table 1. PACE Interactive Access Mode Latency and Packet Loss Rate Bounds

attemptLimitMax.	Access Latency	Packet Loss Rate
6	3.23 ms	0.002 E-3
7	4.83 ms	0.24 E-6
8	8.17 ms	1.86 E-9
9	14.80 ms	7.28 E-12
10	28.00 ms	1.42 E-14
11	54.24 ms	1.39 E-17
12	54.30 ms	6.78 E-21
13	54.37 ms	1.65 E-24

of the attemptLimit. The access latency is computed by adding the worst case collision overhead while in “Trying to Transmit” state, the worst case backoff of the IEEE 802.3 node while in “Wait to Receive” state, and the transmission time of a maximum-length packet. The packet loss rate (PLR) is computed by calculating the probability of reaching attemptLimit in “Trying to Transmit” state (i.e. the probability of IEEE 802.3 node picking 0 as backoff attemptLimit times in a row).

A typical maximum attempt of seven for multimedia applications would give an access latency of around 4.8 ms and a maximum

PLR of 0.24(10⁻⁶ (i.e., one packet in over four million dropped). The bit error rate (BER) is specified at no worse than 1 in 10⁸ for 10BASE-T. Thus, the PLR due to bit errors can be as high as approximately 1 in 173,600 for minimum-size packets and 1 in 8,500 for maximum size packets. Compared to the worst-case PLR as a result of bit errors, the PLR for a maximum of seven attempts by the Interactive Access algorithm is more than an order of magnitude less.

Table 2 shows the interleaving traffic pattern when PACE Interactive Access is enabled at the switch port. Table 3 on page 6 shows the traffic trace of the same link without the benefit of PACE Interactive Access, clearly demonstrating the capture effect.

Figure 1 on page 7 depicts the test configuration for throughput and packet loss under link overload condition. Figures 2 and 3 on page 7 show the throughput and the packet loss test results respectively. The lower bandwidth utilization of 64 byte packets on PACE technology-enabled links reflects overhead resulting from enforcing the interleaved access. For real-world networks this is not an issue, since on average the packet sizes are greater than minimum packet size. The real

Table 2. Traffic Traces of a Switched Ethernet Link with PACE Interactive Access

Sniffer Network Analyzer data from 1-Mar-95 at 14:55:10, file C:\CAPTURE\PACE.ENC, Page 1

	Delta T	Destination	Source	Summary
M 1		006008C00023	3Com AB00EF	DSAP=3C, I frame
2	0.001	3Com ABCDEF	026008C00123	NETB D=07 S=05 Data, 1468 bytes
3	0.0013	006008C00023	3Com AB00EF	DSAP=3C, I frame
4	0.0012	3Com ABCDEF	026008C00123	NETB D=07 S=05 Data, 1468 bytes
5	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
6	0.0012	3Com ABCDEF	026008C00123	NETB D=07 S=05 Data, 1464 bytes
7	0.0013	006008C00023	3Com AB00EF	DSAP=3C, I frame
8	0.0003	3Com ABCDEF	026008C00123	NETB D=07 S=05 Data ACK
9	0.0002	026008C00123	3Com ABCDEF	NETB D=05 S=07 Data ACK
10	0.0009	3Com AB00EF	006008C00023	DSAP=3C, I frame
11	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
12	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
13	0.0013	006008C00023	3Com AB00EF	DSAP=3C, I frame
14	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame

Table 2. Traffic Traces of a Switched Ethernet Link with PACE Interactive Access (cont.)

15	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
16	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
17	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
18	0.0012	3Com ABCDEF	026008C00123	NETB D=07 S=05 Data, 1468 bytes
19	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
20	0.0012	3Com ABCDEF	026008C00123	NETB D=07 S=05 Data, 1468 bytes
21	0.0013	006008C00023	3Com AB00EF	DSAP=3C, I frame
22	0.0012	3Com ABCDEF	026008C00123	NETB D=07 S=05 Data, 1468 bytes
23	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
24	0.0012	3Com ABCDEF	026008C00123	NETB D=07 S=05 Data, 1468 bytes
25	0.0013	006008C00023	3Com AB00EF	DSAP=3C, I frame

Table 3. Traffic Traces of a Switched Ethernet Link Without PACE Interactive Access

Sniffer Network Analyzer data from 1-Mar-95 at 15:05:04, file C:\CAPTURE\ETH.ENC, Page 1

	<u>Delta T</u>	<u>Destination</u>	<u>Source</u>	<u>Summary</u>
M 1		006008C00023	3Com AB00EF	DSAP=3C, I frame
2	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
3	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
4	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
5	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
6	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
7	0.0012	006008C00023	3Com AB00EF	DSAP=3C, I frame
8	0.0013	006008C00023	3Com AB00EF	DSAP=3C, I frame
9	0.0013	006008C00023	3Com AB00EF	DSAP=3C, I frame
10	0.0005	026008C00123	3Com ABCDEF	NETB D=05 S=07 Data, 619 bytes
11	0.0077	3Com AB00EF	006008C00023	DSAP=3C, I frame
12	0.0014	006008C00023	3Com AB00EF	DSAP=3C, I frame
13	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
14	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
15	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
16	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
17	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
18	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
19	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
20	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
21	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
22	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
23	0.0005	3Com ABCDEF	026008C00123	NETB D=07 S=05 Data, 619 bytes
24	0.0012	3Com AB00EF	006008C00023	DSAP=3C, I frame
25	0.0013	006008C00023	3Com AB00EF	DSAP=3C, I frame

significant difference is seen in Figure 3, where the PACE technology-enabled links experienced no or very little packet loss even under the most demanding traffic scenario.

Performance of PACE Interactive Access for 100 Mbps Link Speed

Figure 4 on page 8 depicts the effect of PACE Interactive Access on a Fast Ethernet link. This test was conducted using the NetBench performance test software between 24 PC clients (66 MHz 486s) on 10 Mbps Ethernet ports and a server (90 MHz Pentium) on 100

Mbps Fast Ethernet port of the same SuperStack® II Switch 1000. When PACE Interactive Access is disabled on all of the ports on the SuperStack II Switch 1000, the aggregate throughput on the Fast Ethernet port is similar to that of the reference switch (a Grand Junction FastSwitch). When PACE Interactive Access is enabled on all of the ports on the same SuperStack II Switch 1000, significantly higher aggregate throughput on the Fast Ethernet link was attained much more quickly and was maintained as the message size increased.

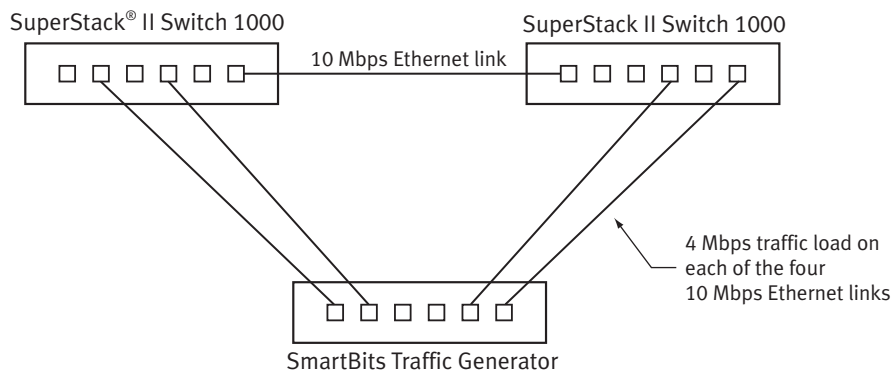


Figure 1. Performance Test Configuration

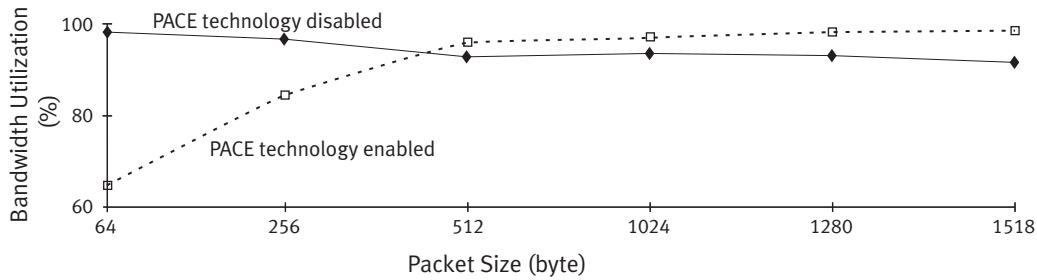


Figure 2. Bandwidth Utilization (Throughput) Comparison

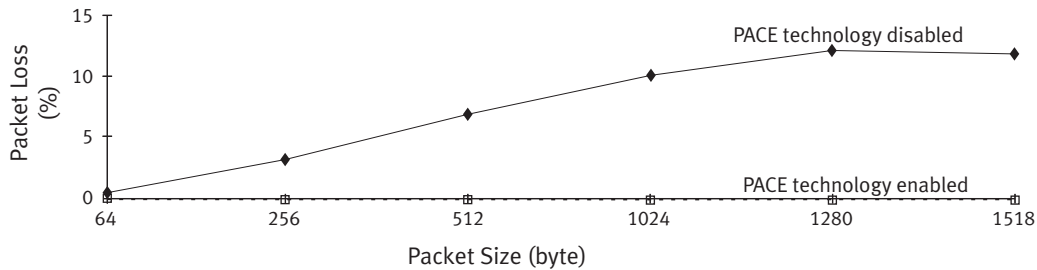


Figure 3. Packet Loss (Throughput) Comparison

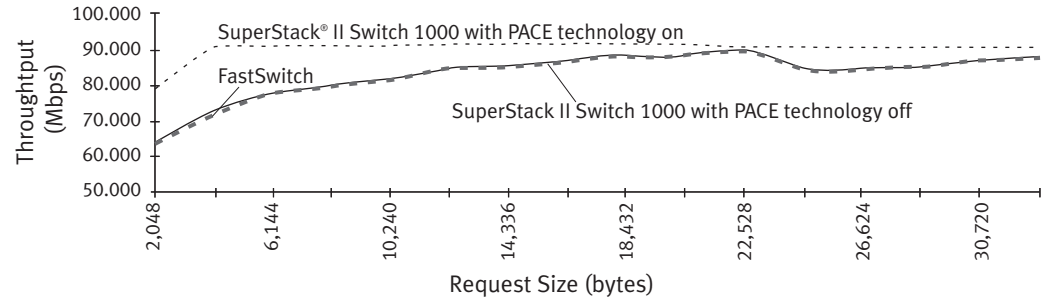


Figure 4. Performance of PACE Interactive Access on Fast Ethernet Link

The test scenario described above parallels that of a video-on-demand application. What was not shown in Figure 4, but most important to multimedia applications, is the well-behaved jitters characteristics provided by PACE Interactive Access. A recent test at the University of Alberta on a similar setup using a real-time video distribution application verifies the projected delay and jitters control characteristics of PACE Interactive Access [16]. These tests clearly show that PACE Interactive Access scales well to the 100 Mbps link speed.

PACE Repeater

While the trend toward switched networks is quite evident, shared LAN segments still have their uses due to their lower cost, particularly in small business and home office environments. Therefore, it is important to provide a low-cost shared LAN technology that can support multimedia applications in a manner compatible with existing Ethernet NICs and wiring. To accomplish this, we have extended the PACE algorithm in the form of PACE Repeater [17]. In this section, we describe how the PACE Repeater algorithm works. We evaluate the performance of PACE Repeater when there are only video streams in the network and show that significantly more throughput can be achieved as compared to regular Ethernet while meeting the end-to-end delay and loss requirements of video traffic. We consider the performance of PACE Repeater when the network traffic consists of bursty data traffic. For this scenario, too, we show that PACE Repeater outperforms Ethernet as long as the packet size is not very

small and the network load is not very high. We also examine the most challenging traffic scenario, video in the presence of bursty data traffic. We show that PACE Repeater can support video traffic in the presence of bursty data traffic far more efficiently as compared to regular Ethernet. Finally, we examine the scalability of PACE Repeater for networks with a large number of active stations, and show that PACE Repeater works well even when the number of active ports is on the order of several tens.

Description of the PACE Repeater Algorithm

In 10BASE-T and 100BASE-T Ethernets, end stations are connected through a multiport repeater, also referred to as a hub, using unshielded twisted pair wires as shown in Figure 5. The repeater's function is simply to forward the electrical signals received on any port to all other ports. Collisions are detected by the end stations when carrier is detected on the incoming direction while the end station is transmitting a packet.

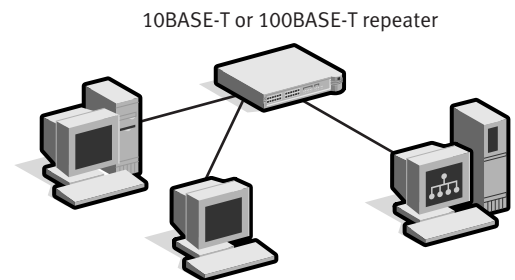


Figure 5. A 10BASE-T or 100BASE-T Ethernet Segment

The PACE Repeater algorithm is based on the idea that the repeater has information about all the stations attempting to transmit packets at a given time, and therefore it can detect collisions and more efficiently resolve them in a centralized manner. The end stations implement the standard 802.3 CSMA/CD protocol without any changes.

When a collision is detected, PACE Repeater selects one of the active stations as the winner. After the end stations go through their backoff period, one of the following three outcomes may occur:

1. The winner station selects a time slot that is earlier than all others; in that case, no collisions would occur and the other stations defer to the winner station until it transmits its packet, just like in regular Ethernet.

2. The winner station and one or more of the other stations select a time slot that is earlier than all others. In that case, the repeater does not forward the other stations' signals to the winner port. The winner station is able to continue transmitting its packet, which gets buffered by the repeater. At the same time, the repeater tries to aggressively push the winner packet through the network by repeatedly attempting to transmit it, backing off in case of a collision only by the Ethernet IPG time, just like in the case of switched PACE Interactive Access algorithm.

3. One or more of the stations may pick a time slot earlier than the one that the winner station picks; in this case, the repeater simply generates artificial collisions to prevent those stations from successfully transmitting their packet until the selected winner station starts transmitting its packet.

For the second and third outcomes, the winner station is likely to transfer its entire packet into the repeater before the repeater completes sending it to the other end stations. In that case, whenever the repeater receives a complete packet from the winner, it sends a preamble signal on the winner port to prevent the winner station from starting to send a second packet.

The repeater can select the winner station according to various different algorithms, the criteria being fairness among the stations, low

latency and jitters, high throughput, etc. One possibility is to have a round-robin scheme among the stations that have packets to transmit. This prevents the capture effect and provides fair access to the network. Another possibility is first come, first served (FCFS) selection, where during a collision the repeater chooses as winner the first port where carrier is detected, with the exception of the port that has transmitted the previous successful packet (to prevent capture).

Performance of PACE Repeater for Video Traffic Alone

We performed computer simulations to measure the performance of PACE Repeater. In the experiments described in this subsection, we used real video traffic generated by encoding one minute of a motion picture (Star Trek VI) using H.261, and controlling the encoder according to constant bit rate buffer-based feedback to achieve a bit rate of 1.536 Mbps [18]. The video streams were packetized using full-size Ethernet packets. We simulated three to six video streams being sent on a 10BASE-T segment (for both regular Ethernet and PACE Repeater), corresponding to a total offered traffic rate from about 4.5 to 9 Mbps. We computed the delay that each packet experienced due to the sum of packet formation (i.e., the time it takes for the encoder to generate sufficient number of bits to fill a packet), queuing, network access, transmission, and propagation. (The end-to-end system model we have used is identical to the one presented in [5].) We show here the results for the round-robin selection of the winner for PACE Repeater, but the results are nearly identical for FCFS as well.

Figure 6 on page 10 shows the resulting average delay as a function of the number of video streams on the network. For up to five video streams and for both standard Ethernet and PACE Repeater, the network delay is negligible and the dominant delay component is due to the packet formation time (which is around 7 to 8 ms). However, when six video streams are sent over the network, the average delay for PACE Repeater still remains the same, while the average delay for standard

Ethernet becomes very large. Similarly, the standard deviation of delay (shown in Figure 7) is about the same for the two networks for up to five video streams, but it becomes very large for standard Ethernet when six streams are sent over the network. This suggests that PACE Repeater can achieve a significantly

higher throughput than regular Ethernet while keeping delay and jitter at low levels.

For video traffic, the average and standard deviation of delay are not sufficient to measure the performance; since packets exceeding a certain delay value are useless, one has to consider the tail distribution of delay as well.

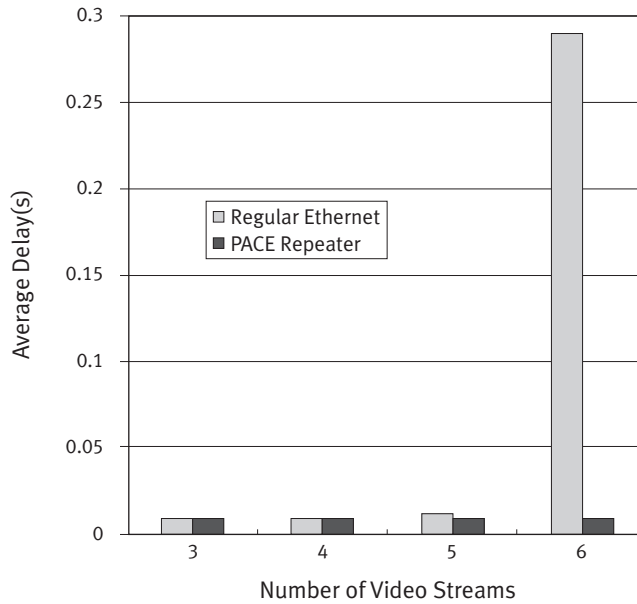


Figure 6. Average Delay vs. Number of Video Streams (each stream has a rate of 1.536 Mbps)

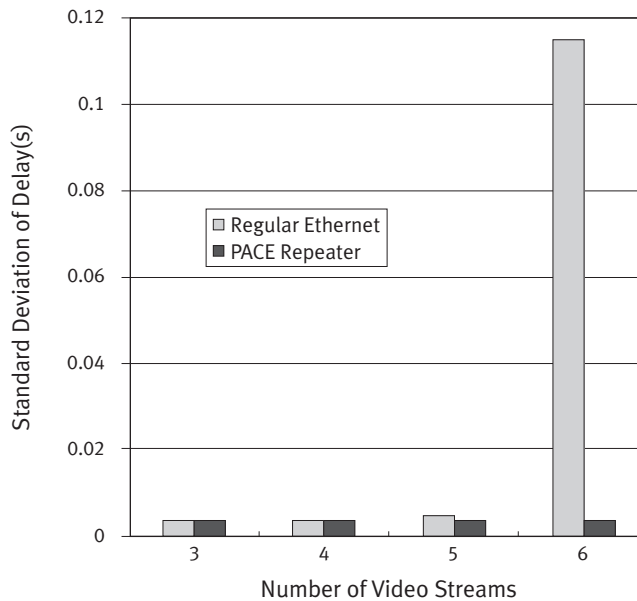


Figure 7. Standard Deviation of Delay vs. Number of Video Streams (each stream has a rate of 1.536 Mbps)

Figure 8 shows the ratio of packets exceeding a given maximum delay constraint (and hence considered lost) as a function of the delay constraint value when there are five video streams on the network. Even though for five video streams the average and standard deviation of the delay values are very close for the two types of networks, the tail distribution differs significantly in favor of PACE Repeater. Indeed, if the maximum delay constraint is, say, 50 ms, then the video streams sent over PACE Repeater experience no packet loss due to exceeding the delay value, while those sent over regular Ethernet experience over 0.1 percent packet loss, which can impair the displayed video quality significantly. Note the flat packet loss region for standard Ethernet when the delay values are over 0.1 seconds. This is the result of packets dropped at the sender when the maximum number of collisions allowed by the Ethernet MAC protocol (i.e., 16 collisions) occurs.

We also compared the performance of standard Ethernet and PACE Repeater at 100 Mbps. Figure 9 shows a graph equivalent to Figure 8, but for 100 Mbps and with 50 active video sources. In this scenario, PACE Repeater also outperforms standard Ethernet,

particularly when the delay constraint is greater than about 30 ms.

Performance of PACE Repeater for Data Traffic Alone

This subsection presents simulation results when the network carries data traffic generated according to a simple two-parameter model that captures the burstiness and total data load. The model consists of fixed-size

messages generated with a uniform inter-arrival time. If a message is greater than the full-size Ethernet packet (i.e., 1500 bytes), it is divided into as many full-size packets as needed, and those packets are queued as a bulk arrival for transmission. We have fixed the number of active stations at five, with the traffic divided equally among them. We have measured the packet delay due to the sum of queuing, network access, transmission, and

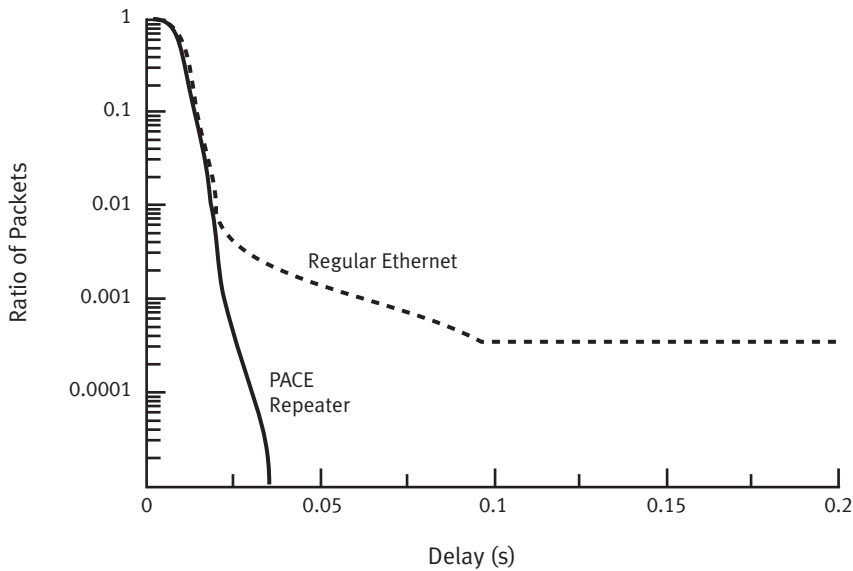


Figure 8. Video Delay Distribution for Five Video Streams of 1.536 Mbps Each

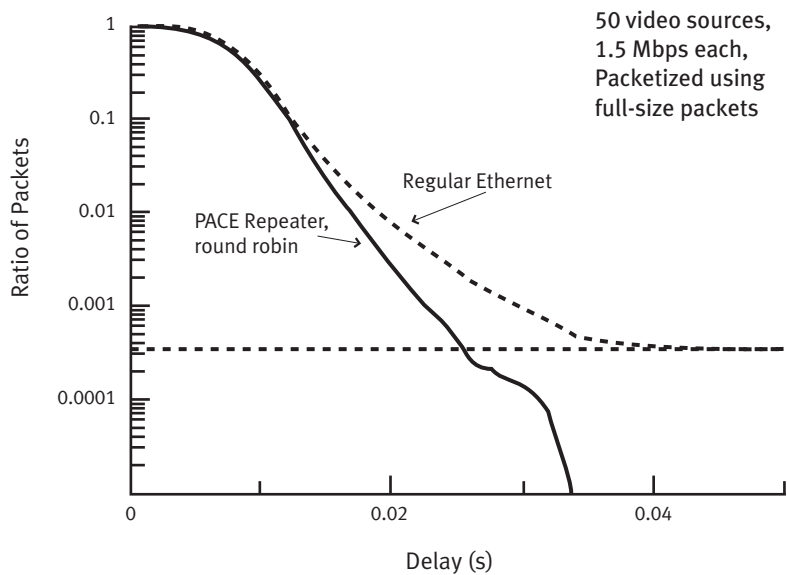


Figure 9. Video Delay Distribution for 100 Mbps Ethernet, 50 Video Streams of 1.536 Mbps Each

propagation delays. Both round robin and FCFS selection methods for PACE Repeater were considered.

Figure 10 shows average packet delay as a function of the data message size, both for regular Ethernet and for PACE Repeater for a total data load of 7 Mbps. It is interesting to note that the dominant component of the delay for PACE Repeater is due to transmission time and queuing of the packets (the latter increasing linearly as the message size

increases, due to the transmission time of the packets queued in front of a given packet). By contrast, standard Ethernet has an extra component of delay due to the capture effect and the associated access overhead, which starts taking place as soon as there are more than one or two packets per message.

A similar result is observed for the standard deviation of delay as well, as shown in Figure 11. This indicates that PACE Repeater helps reduce the access time not only for the

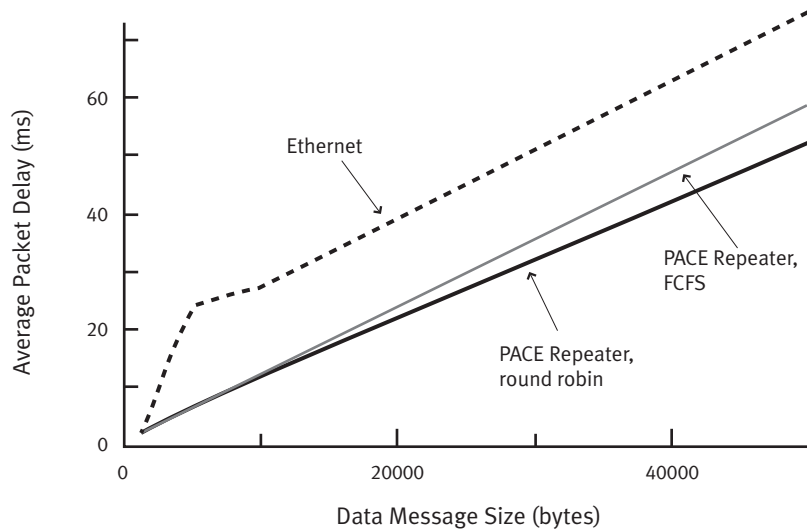


Figure 10. Average Packet Delay vs. Data Message Size (total data load=7 Mbps)

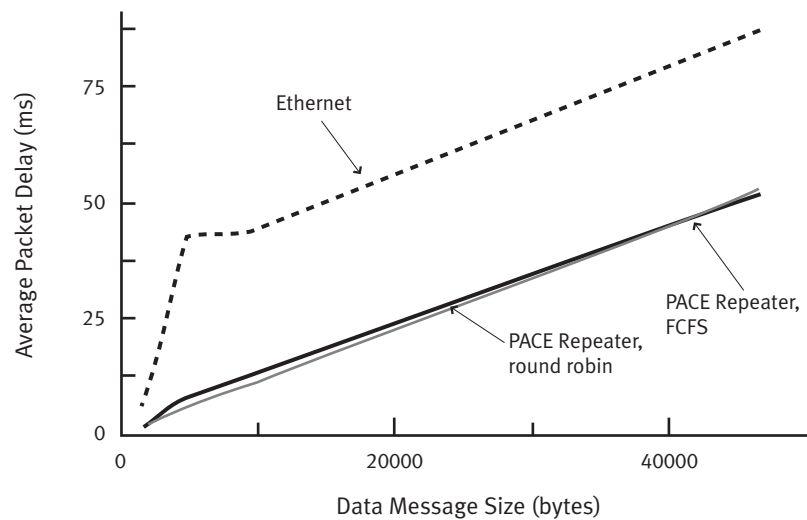


Figure 11. Standard Deviation of Packet Delay vs. Data Message Size (total data load=7 Mbps)

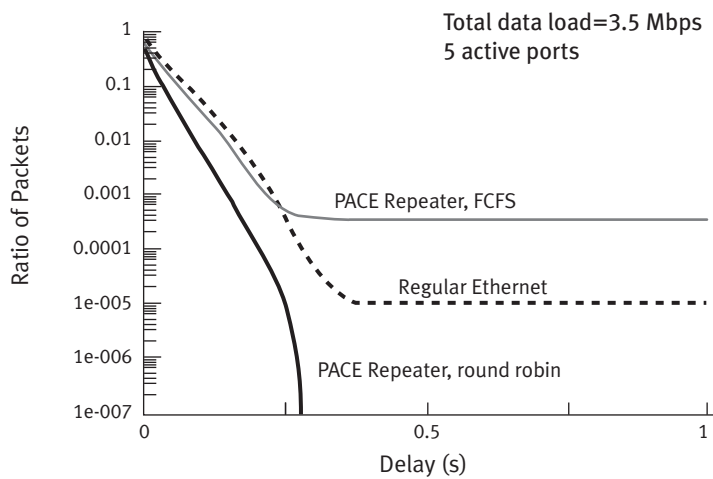
well-behaved audio/video streams, but also for bursty data transmissions such as file transfers, Web access, file server access, etc. It is also interesting to note that the performance of round-robin and FCFS schemes are quite similar.

As a benchmark test similar to the one discussed earlier for PACE Interactive Access, we have simulated the network for very small packets. We have chosen the payload size to be 46 bytes (with 18 bytes of MAC header and CRC, this corresponds to the smallest possible Ethernet packet), and simulated the network with five active stations. Figure 12 shows the ratio of packets exceeding a given delay

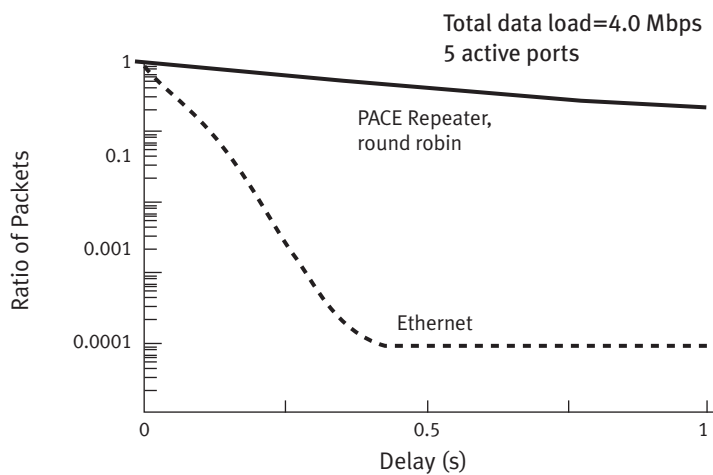
constraint as a function of the delay for such a scenario.

In part (a), for a total data load equal to 3.5 Mbps, it is interesting to note that the regular Ethernet outperforms PACE Repeater with FCFS selection, but is not as good as PACE Repeater with round-robin selection. This is because round-robin selection allows fair access to all ports, while FCFS may at times favor some stations over others.

In part (b), for a total data load equal to 4 Mbps (not including the headers), one can see that standard Ethernet outperforms even round-robin PACE Repeater when the data load is increased to 4 Mbps. The reason is



(a) Total Data Load = 3.5 Mbps



(b) Total Data Load = 4 Mbps

Figure 12. Benchmark Tests with Minimum-Size Packets

similar to that explained in the section on Interactive Access (i.e., interleaving the network access for very small packets introduces a high overhead). Of course, the condition of a high traffic load comprised mainly of very small packets is not a very realistic one. For the more realistic scenarios of either a high load comprised of large packets, or a low load comprised of small packets, we find that PACE Repeater outperforms Ethernet.

Performance of PACE Repeater for Integrated Video and Data Services

Considering the performance results for video alone and data alone, one would expect that PACE Repeater should support video well when the two traffic types are mixed together, and this is confirmed by simulation results. As an example, Figure 13 plots the ratio of packets exceeding a given delay constraint as a function of the delay constraint for one video source generating 1.536 Mbps of traffic, in the presence of three data sources generating a total of 5.4 Mbps of traffic with a message size of 10 KB.

Figure 13 clearly shows that standard Ethernet suffers heavily under such a traffic scenario, losing around 0.1 percent of video packets even for very large delay constraints such as 400 ms. By contrast, PACE Repeater (in both varieties) performs very well, incurring negligible packet loss beyond a delay constraint of 100 ms.

PACE Repeater Scalability

For a given traffic load, it is well known that the performance of CSMA/CD decreases as the number of stations generating the traffic is increased. The reason is that at a given time, the probability of finding more than one station trying to transmit and colliding is higher for a greater number of active nodes. This is especially true at relatively high loads, where momentary backlogs can occur, increasing the likelihood that multiple deferring stations will collide at the end of a packet transmission.

One logically expects the performance to decrease as the number of active ports is increased for PACE Repeater as well. We conducted simulations to quantify this effect. In particular, we simulated the standard Ethernet and PACE Repeater for 5 and 25 active ports, using the same two-parameter data traffic model described previously. The data message size was set to be equal to 1500 bytes, and both round-robin and FCFS selection algorithms were considered.

Figure 14 shows the average delay as a function of the data load. It can be clearly seen that the average delay for Ethernet is significantly greater for 25 ports than for 5 ports. For PACE Repeater there is still some difference between 5 and 25 port scenarios, but it is not as pronounced.

Figure 15 shows the standard deviation of delay as a function of the data load for the

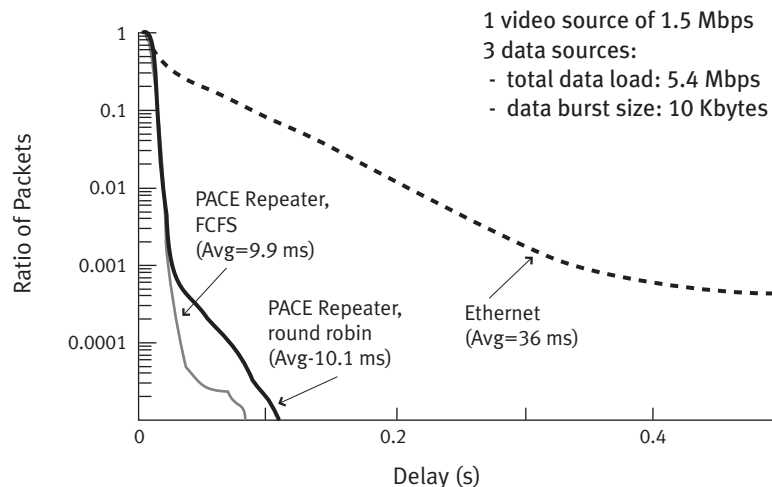


Figure 13. Video Delay Distribution in the Presence of Bursty Data Traffic

same scenario. Here, the situation is somewhat different. The difference in standard deviation between 5 and 25 ports for PACE Repeater is around the same order of magnitude as that of the standard Ethernet. Nevertheless, it is significant to note that, as far as both average and standard deviation of delay are concerned, the performance of PACE Repeater does not decrease very much from 5 to 25 ports. In both cases, PACE Repeater remains signifi-

cantly better than the performance of standard Ethernet.

Further simulation at 100 ports verified that the performance of PACE Repeater remains good even for such a large number of ports. Similar test cases for 100 Mbps link speeds with 50 and 100 active ports again confirmed the same performance scaling characteristics.

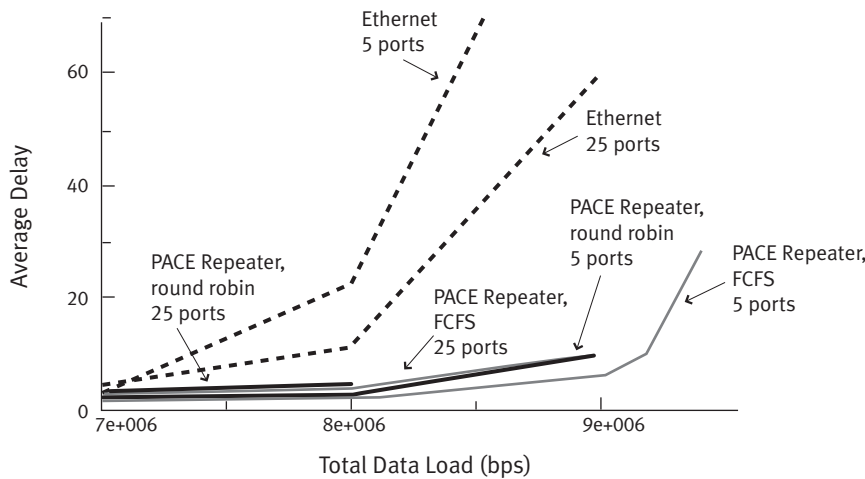


Figure 14. Average Delay vs. Data Load for Ethernet and PACE Repeater for 5 and 25 Active Ports

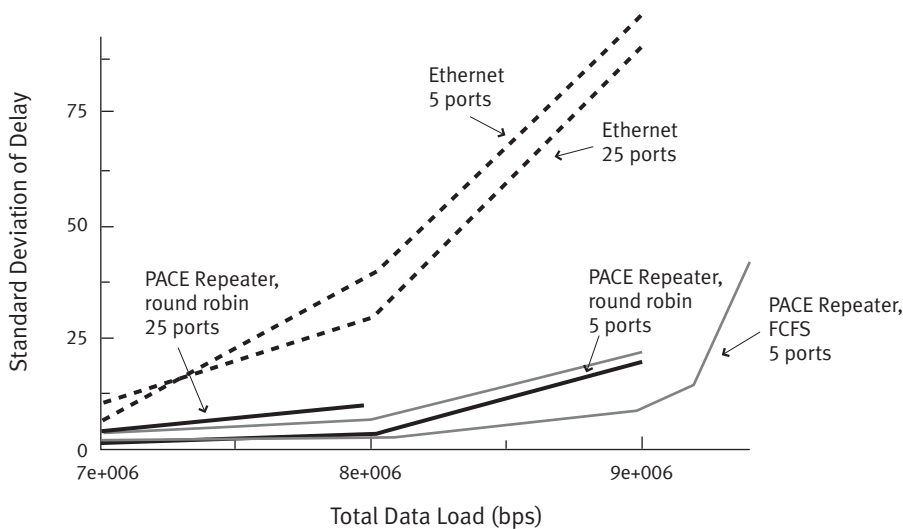


Figure 15. Standard Deviation of Delay vs. Data Load for Ethernet and PACE Repeater for 5 and 25 Active Ports

Conclusion

This paper introduced PACE Interactive Access technology for switched Ethernets, as well as a generalization of the PACE Interactive Access algorithm for a shared LAN segment, PACE Repeater. Both have been devised as enhancements to the Ethernet MAC protocol with the goal of supporting real-time multimedia traffic.

Bench tests demonstrated the ability of the PACE Interactive Access scheme to provide significantly improved throughput, latency, jitter, and packet loss characteristics as compared to the standard Ethernet. Computer simulations showed that PACE Repeater can successfully support multimedia traffic, including mixtures of video and bursty data traffic, and is scalable to large shared segments, with the number of active ports on the order of several tens. These results are in contrast with the standard Ethernet, where the video streams break down in the presence of bursty data traffic even under light to moderate network utilization.

As enterprises migrate from shared to switched network infrastructures, the migration of desktops to full-duplex Ethernet and Fast Ethernet will occur at a significantly slower rate. PACE technology-enabled workgroup switches satisfy the pent-up demand for real-time, networked multimedia applications by allowing the deployment of such applications to desktops over existing half-duplex links.

Today there is rapid growth of Ethernet-connected PCs in small offices. The convergence of computers and consumer electronics will drive similar adoptions within homes in the near future. Various types of real-time multimedia applications, such as IP telephony and multi-party games, will join current mainstream applications within these environments. PACE Repeater-enabled Ethernet and Fast Ethernet hubs offer a low-cost multimedia networking alternative suitable for small offices and homes. ◻

References

1. ISO/IEC 8802-3: 1993 [ANSI/IEEE Std. 802.3, 1993], Information technology - Local and metropolitan area networks - Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications.
2. G.J. Nutt and D.L. Bayer, "Performance of CSMA/CD Networks Under Combined Voice and Data Loads," *IEEE Transactions on Communications*, vol. COM-30, no. 1, pp. 6-11, January 1982.
3. I. Chlamtac and M. Eisinger, "Performance of Integrated Services (Voice/Data) on CSMA/CD Networks," *Proceedings of ACM Sigmetrics '85*, pp. 87-93, August 1985.
4. T.A. Gonsalves, "Comparative Performance of Broadcast Bus Local Area Networks with Voice and Data Traffic," Stanford University Technical Report CSL:TR-87-137, March 1987.
5. F.A. Tobagi and I. Dalgıç, "Evaluation of 10BASE-T and 100BASE-T Ethernets Carrying Constant Bit Rate Video Traffic," *IEEE JSAC*, special issue on Distributed Multimedia Systems and Technology, v14, no 7, pp. 1436-1454, September 1996.
6. S. Tasaka, "Dynamic Behavior of a CSMA-CD System with a Finite Population of Buffered Users," *IEEE Transactions on Communications*, vol. COM-34, no. 6, pp. 576-586, June 1986.
7. N. F. Maxemchuk, "A Variation on CSMA/CD that Yields Movable TDM Slots in Integrated Voice/Data Local Networks," *Bell System Technical Journal*, vol. 61, no. 7, pp. 1527-1550, September 1982.
8. M. Molle, "Binary Logarithmic Access Method," *IEEE 802.3w*.
9. E. Arthurs and B. W. Stuck, "A Modified Access Policy for ETHERNET Version 1.0 Data Link Layer," *IEEE Transactions on Communications*, vol. COM-32, no. 8, pp. 977-979, August 1984.
10. M. Fine and F. A. Tobagi, "Demand Assignment Multiple Access Schemes in Broadcast Bus Local Area Networks," *IEEE Transactions on Computers*, vol. C-33, no. 12, pp. 1130-1159, December 1984.
11. G. L. Choudhury and S. S. Rappaport, "Priority Access Schemes Using CSMA-CD," *IEEE Transactions on Communications*, vol. COM-33, no. 7, pp.620-626, July 1985.
12. M. Fine and F. A. Tobagi, "Packet Voice on a Local Area network with Round Robin Service," *IEEE Transactions on Communications*, vol. COM-34, no. 9, pp. 906-915, September 1986.
13. S. M. Sharrock, S. Ghanta, H. C. Du and K. J. Maly, "CSMA/CD-based, integrated Voice/Data Protocol with Dynamic Channel Allocation," *Computer Networks: The International Journal of Distributed Informatique*, vol. 18, no. 1, pp. 1-18, November 1989.
14. W.P. Sherer, L.C. Lo and J. Hickey, "Method and Apparatus for Controlling Latency and Jitter in a Local Area Network Which uses a CSMA/CD Protocol," U.S. Patent No. 5,568,469.
15. 3Com Ref:426-PACE-1.04: PACE Ethernet Interactive Access Algorithm, March 1996.
16. B. Chen, "Test Results of 3Com PACE Technology as Tested at the University of Alberta," University of Alberta, Canada, September 1996. (Available at <http://www.ualberta.ca/~nos/pace/3com-report.htm>)
17. W.P. Sherer, W.T. Tang, I. Dalgic, P. Wang, "Method and Apparatus for Controlling Latency and Jitter in Shared CSMA/CD (Repeater) Environment," U.S. Patent Application pending.
18. I. Dalgıç and F.A. Tobagi, "Characterization of Quality and Traffic for Various Video Encoding Schemes and Various Encoder Control Schemes," Stanford University Technical Report CSL-TR-96-701, August 1996.



More connected.™

3Com Corporation enables individuals and organizations worldwide to stay more connected by communicating and sharing information and resources anytime, anywhere. As one of the world's preeminent suppliers of data, voice, and video communications technology, 3Com has delivered networking solutions to more than 200 million customers worldwide. The company provides large enterprises, small and medium enterprises, carriers and network service providers, and consumers with comprehensive, innovative information access products and system solutions for building intelligent, reliable, and high-performance local and wide area networks.

3Com Corporation

P.O. Box 58145
5400 Bayfront Plaza
Santa Clara, CA
95052-8145
Phone: 1 800 NET 3Com
or 1 408 326 5000
Fax: 1 408 326 5001
World Wide Web:
www.3com.com

Asia Pacific Rim

Sydney, Australia
Phone: 61 2 9937 5000
Fax: 61 2 9956 6247

Melbourne, Australia
Phone: 61 3 9934 8888
Fax: 61 3 9934 8880

Beijing, China
Phone: 86 10 68492 568
Fax: 86 10 68492 789

Shanghai, China
Phone: 86 21 6350 1581
Fax: 86 21 6350 1531

Hong Kong
Phone: 852 2501 1111
Fax: 852 2537 1149

India
Phone: 91 11 644 3974
Fax: 91 11 623 3192

Indonesia
Phone: 62 21 572 2088
Fax: 62 21 572 2089

Osaka, Japan
Phone: 81 6 536 3303
Fax: 81 6 536 3304

Tokyo, Japan
Phone: 81 3 5977 3266
Phone: 0120 313266
(toll free from Japan)
Fax: 81 3 5977 3370

Korea
Phone: 82 2 3455 6300
Fax: 82 2 319 4710

Malaysia
Phone: 60 3 715 1333
Fax: 60 3 715 2333

New Zealand
Phone: 64 9 366 9138
Fax: 64 9 366 9139

Philippines
Phone: 632 892 4476
Fax: 632 811 5493

Singapore

Phone: 65 538 9368
Fax: 65 538 9369

Taiwan

Phone: 886 2 2 377 5850
Fax: 886 2 2 377 5860

Thailand

Phone: 662 231 8151 5
Fax: 662 231 8158

3Com Austria

Phone: 43 1 580 17 0
Fax: 43 1 580 17 20

3Com Benelux B.V.

Belgium
Phone: 32 2 725 0202
Fax: 32 2 720 1211

Netherlands
Phone: 31 346 58 62 11
Fax: 31 346 58 62 22

3Com Canada

Calgary
Phone: 1 403 265 3266
Fax: 1 403 265 3268

Edmonton
Phone: 1 403 423 3266
Fax: 1 403 423 2368

Montreal
Phone: 1 514 683 3266
Fax: 1 514 683 5122

Ottawa
Phone: 1 613 566 7055
Fax: 1 613 233 9527

Toronto
Phone: 1 416 498 3266
Fax: 1 416 498 1262

Vancouver
Phone: 1 604 434 3266
Fax: 1 604 434 3264

3Com Eastern Europe/CIS

Bulgaria
Phone: 359 2 962 5222
Fax: 359 2 962 4322

Czech/Slovak Republics
Phone: 420 2 21845 800
Fax: 420 2 21845 811

Hungary
Phone: 36 1 250 8341
Fax: 36 1 250 8347

Poland
Phone: 48 22 6451351
Fax: 48 22 6451352

Russia

Phone: 7 095 258 09 40
Fax: 7 095 258 09 41

3Com France

Phone: 33 1 69 86 68 00
Fax: 33 1 69 07 11 54
Carrier and Client Access
Phone: 33 1 41 97 46 00
Fax: 33 1 49 07 03 43

3Com GmbH

Berlin, Germany
Phone: 49 30 3498790
Fax: 49 30 34987999

Munich, Germany
Phone: 49 89 627320
Fax: 49 89 62732233

3Com Iberia

Portugal
Phone: 351 1 3404505
Fax: 351 1 3404575

Spain
Phone: 34 1 509 69 00
Fax: 34 1 307 79 82

3Com Latin America

U.S. Headquarters
Phone: 1 408 326 2093
Fax: 1 408 764 5730

Miami, Florida
Phone: 1 305 461 8400
Fax: 1 305 461 8401

Argentina
Phone: 54 1 312 3266
Fax: 54 1 314 3329

Brazil
Phone: 55 11 246 5001
Fax: 55 11 246 3444

Chile (serving Chile, Bolivia, and Peru)
Phone: 562 240 6200
Fax: 562 6240 6231

Colombia
Phone: 57 1 629 4847
Fax: 57 1 629 4503

Mexico
Phone: 52 5 520 7841/7847
Fax: 52 5 520 7837

Peru
Phone: 51 1 221 5399
Fax: 51 1 221 5499

Venezuela

Phone: 58 2 953 8122
Fax: 58 2 953 9686

3Com Mediterraneo

Milan, Italy
Phone: 39 2 253011
Fax: 39 2 27304244

Rome, Italy
Phone: 39 6 5279941
Fax: 39 6 52799423

3Com Middle East

Phone: 971 4 319533
Fax: 971 4 316766

3Com Nordic AB

Denmark
Phone: 45 48 10 50 00
Fax: 45 48 10 50 50

Finland
Phone: 358 9 435 420 67
Fax: 358 9 455 51 66

Norway
Phone: 47 22 58 47 00
Fax: 47 22 58 47 01

Sweden
Phone: 46 8 587 05 600
Fax: 46 8 587 05 601

3Com Southern Africa

Phone: 27 11 807 4397
Fax: 27 11 803 7405

3Com Switzerland

Phone: 41 844 833 933
Fax: 41 844 833 934

3Com UK Ltd.

Edinburgh
Phone: 44 131 240 2900
Fax: 44 131 240 2903

Ireland
Phone: 353 1 820 7077
Fax: 353 1 820 7101

Manchester
Phone: 44 161 873 7717
Fax: 44 161 873 8053

Marlow
Phone: 44 1628 897000
Fax: 44 1628 897003

To learn more about 3Com products and services, visit our World Wide Web site at www.3com.com. 3Com is a publicly traded corporation (Nasdaq: COMS).

Copyright © 1998 3Com Corporation. All rights reserved. 3Com, the 3Com logo and SuperStack are registered trademarks of 3Com Corporation. More connected. and PACE are trademarks of 3Com Corporation. Pentium is a trademark of Intel Corporation. Other product and brand names may be trademarks or registered trademarks of their respective owners. All specifications are subject to change without notice.



Printed in U.S.A. on recycled paper

500676-001 9/98